EXHIBIT F

IIIORRISON FOERSTER

707 WILSHIRE BOULEVARD
SUITE 6000
LOS ANGELES
CALIFORNIA 90017-3543

TELEPHONE: 213.892.5200
FACSIMILE: 213.892.5454

WWW.MOFO.COM

MORRISON & FOERSTER LLP

AUSTIN, BEIJING, BERLIN, BOSTON,
BRUSSELS, DENVER, HONG KONG,
LONDON, LOS ANGELES, MIAMI,
NEW YORK, PALO ALTO, SAN DIEGO,
SAN FRANCISCO, SHANGHAI, SINGAPORE,
TOKYO, WASHINGTON, D.C.

August 20, 2024

Zach Savage
Susman Godfrey L.L.P.
zsavage@susmangodfrey.com
One Manhattan West
New York, New York 10001

Writer's Direct Contact
+1 (213) 892-5656
ABennett@mofo.com

Re: *The New York Times Company v. Microsoft Corporation, et al.,*
      Case No.: 23-cv-11195-SHS

Counsel,

Please see below for our responses to your letter regarding OpenAI's Responses and Objections to The Times's Second Set of Requests for Production of Documents.

1.   Global Issues

a.   OpenAI's Requests to Meet and Confer

OpenAI did not "merely offer[] to meet and confer"– we offered specific objections to specific RFPs as to why we cannot commit to producing documents in response to them. In particular, OpenAI objected to most of the RFPs identified by The Times in its letter as overbroad, unduly burdensome, and vague. When The Times ignores those objections and suggests that the only thing that matters is relevance, it ignores "both the letter and spirit of the Federal Rules." Fed. R. Civ. P. Rule 26(b)(1).

Moreover, as OpenAI's responses made clear, OpenAI has already agreed to produce documents falling into several of the categories it identifies in its letter.

- The Training Datasets for the Text Models Used to Power ChatGPT: OpenAI has agreed to make these datasets available for inspection (*see, e.g.*, RFP 1) and has either produced or agreed to produce documents related to those datasets through custodial and non-custodial searches. (RFPs 6, 17, 20, 23, 24, 32, 34)

- The Functioning of OpenAI's Products: OpenAI has agreed to produce documents reflecting the functionality and technical documentation of OpenAI products. (RFPs 12, 13, 28, 30, 31)

- User Engagement with OpenAI's Products: OpenAI has already produced documents responsive to RFP 42 (seeking documents concerning OpenAI's "efforts and abilities to monitor, log, analyze, quantify, assess, evaluate, or review user queries within any

IIIORRISON FOERSTER

August 20, 2024
Page Two

> [of OpenAI's products]") and RFP 43 seeking "Any surveys, studies, analyses, or statistics on patterns and practices of [OpenAI's products]").

- OpenAI Financial Information: OpenAI has already produced its profit and loss statement for ChatGPT. (RFP 70)

We requested to meet and confer so that we can better understand what else you're seeking beyond what we have already produced, and why you believe producing that information would be proportionate to the needs of the case. By doing so we can hopefully efficiently resolve any disputes over these requests. We look forward to doing so.

b. The Products and Models at Issue

OpenAI disagrees that its willingness to resolve an earlier dispute regarding the "Generative AI Products and Services" is relevant here. OpenAI expressly withdrew the "analogous limitation" without conceding relevance in order to resolve that dispute. That said, OpenAI is open to reaching a similar compromise to resolve a dispute here as well. Your letter is not clear as to the relevant RFPs at issue. It identifies five RFPs as examples ("see, e.g."), one of which (RFP No. 5) is not obviously relevant to the issue; another (RFP No. 36) is not related to the "Generative AI Products and Services," but is focused on the "Generative AI Models." We are open to discussing whether we would agree to include Bing Chat and Copilot (a product that is not OpenAI's) within the scope of our response to certain RFPs to resolve any dispute, but we require more clarity on the RFPs to which The Times's request applies.

Additionally, if the Court rules in Plaintiff's favor with respect to what models are at issue, OpenAI will of course evaluate how and whether the order applies to these RFPs.

c. Documents Sufficient to Show The Requested Information

OpenAI's "sufficient to show" limitations are not improper. Each of The Times's requests at issue use the precise sort of language targeting "general categor[ies] or group[s] of documents" that courts deem "unduly burdensome on its face." *Laryngeal Mask Co. Ltd. v. Ambu A/S*, No. 07CV1988-DMS(NLS), 2009 WL 10672487, at *2 (S.D. Cal. Jan. 27, 2009). The case cited by The Times does not stand for the general proposition that a party may not respond with a "sufficient to show" limitation; it only holds that a single RFP targeting a specific category of documents that is "central to Plaintiffs' theory of the case" may not be subject to a "sufficient to show" limitation. *In re Stubhub Refund Litig.*, No. 20MD02951HSGTSH, 2022 WL 1640304, at *2 (N.D. Cal. May 24, 2022). The Times makes no effort to explain why any of the RFPs that it identifies meet that bar. Is The Times's position that all of the RFPs meet that bar? We note that RFP Nos. 26 and 48 do not even include the limitation at issue. If The Times believes that any of the requests meet the standard set in *StubHub*, OpenAI is willing to meet and confer to discuss that position.

2. Specific RFPs

⌐⌐ORRISON ⌐OERSTER

August 20, 2024
Page Three

   a. RFPs 53, 54, 55, 61, 65, 66, 69

   RFP 53: Thank you for explaining why you believe that this request is relevant. OpenAI has already agreed to produce "responsive, non-privileged documents in its possession, custody, or control, if any, sufficient to show ChatGPT's approach to paywalls that it locates pursuant to a reasonable and diligent search." (RFP No. 44.) For the avoidance of doubt, OpenAI will confirm that as part of that production, it will conduct a reasonable search for and produce responsive, non-privileged documents, if any, related to the Browse feature's interactions with paywalls and the role, if any, that feature played in the decision to suspend Browse with Bing and/or Bing Chat, to the extent such documents exist.

   RFP 54: OpenAI agrees to produce non-privileged documents related to its statement that "the Times paid someone to hack OpenAI's products." OpenAI notes, however, that The Times is already in possession of many, if not all, of those documents—specifically the prompt-output pairs that OpenAI asked The Times to produce, which The Times has refused to do. (ECF No. 124.) The Times should provide these responsive documents to OpenAI. Moreover, the response is plainly overbroad; as written, it would seem to call for work product and privileged communications, including, for example, drafts of OpenAI's motion to dismiss.

   RFP 55: OpenAI maintains its objection that this RFP is not relevant to the claims and defenses in this litigation. The article you cite is of no help. "AI safety and ethics" are only relevant, if at all, to the extent that they implicate the copyright or trademark issues in this case. The Times has not met its burden to show that these issues are relevant to the departures of any of OpenAI's former employees. Moreover, this request, even if limited to copyright or trademark issues, is disproportionate to the needs of this case. The Times would have OpenAI engage in an extensive search of highly sensitive employee-related documents on the off-chance copyright or trademark related issues are mentioned. OpenAI will not invade the privacy of its current and former employees for documents of no, or at best marginal relevance. The Times has failed to show a need for this information, especially in light of these concerns. Additionally, OpenAI interprets the RFP to be seeking "[d]ocuments concerning any . . . disciplinary actions, departures . . . or personnel-related actions Defendants have taken or considered concerning individuals or teams involved with ethics, harm reduction, safety, lawfulness, and responsible AI initiatives." (emphasis added.) It was on that basis that we objected to it as vague and ambiguous. OpenAI will not permit The Times to engage in a fishing expedition. If Plaintiffs are limiting the scope of the RFP to the specific employees and board members to whom this request applies, OpenAI will consider that.

   RFP 61: OpenAI understands the requested documents to concern OpenAI's potential legal liabilities. Such documents are likely to be privileged legal communications. It would be unduly burdensome and disproportionate to the needs of the case to search for and review these communications on the off-chance there is some non-privileged discussion. Moreover, OpenAI does not understand from The Times's letter how the requested

IIIORRISON FOERSTER

August 20, 2024
Page Four

documents are relevant to its infringement claims. A conclusory assertion of relevance does not meet The Times's burden. Lastly, OpenAI directs The Times to OPCO_NYT_0000371, which contains certain responsive, non-privileged information.

RFP 65: We are confused by this portion of your letter. Are you seeking public relations strategies or marketing strategies? These are two different fields. We understood your request to be related to the former and refused to produce documents on that basis, including, at least, because such a request is overbroad. If The Times is now asserting that it seeks marketing documents, we can take that back to our client. Please clarify what documents The Times is seeking with this RFP.

Also, we are not clear on your relevance claims. Can you explain why specifically these documents would be relevant to "willfulness, fair use and the substitutive effect of Defendants' products on journalism." We would like to better understand your position before fully responding.

RFP 66: The request for "documents concerning fundraising efforts" is overly broad and unduly burdensome. As written, the request would, for example, appear to encompass almost any email to any potential investor. Given how few such emails are likely to be relevant, the request is not proportional to the needs of the case. This is particularly true when The Times says it's seeking financial information to quantify OpenAI's profits considering that OpenAI has already agreed to produce such information in response to RFP No. 71. OpenAI is willing to meet and confer to better understand why The Times believes documents specifically related to fundraising efforts are relevant and not duplicative with the documents OpenAI has agreed to produce in response to RFP No. 71.

RFP 69: OpenAI objected to the extent any of the information or documents sought by the RFP are publicly available. At least the following responsive documents are publicly available: https://openai.com/charter/; https://openai.com/our-structure/.

In addition, you have not explained why these documents are relevant, other than stating, without any elaboration, that they relate to fair use. Can you explain why documents regarding public access to OpenAI governing documents, financial statements, and conflict of interest rules are relevant to fair use? We do not understand your assertion.

b. RFPs 17, 20, 21, 22, 23, 26, 27, 28, 31, 32, 36, 43, 44, 49, 50, 60, 62, 70, 71

RFP 17: OpenAI can confirm that it has produced documents responsive to this request, including the agreements related to text training data used to train models that have powered ChatGPT. OpenAI objected to this RFP as overly broad, unduly burdensome and disproportionate to the needs of the case because the categories of communications related to training data—an extremely wide swathe of material—go well beyond what is plausibly relevant. For that reason, OpenAI offered to meet and confer regarding an appropriately narrow scope for the portions of the request for which we have not already produced. Please

IIIORRISON FOERSTER

August 20, 2024
Page Five

be prepared to define the universe of communications you're looking for, including the types of training data for which you believe communications would be relevant, and where you believe there is a gap in OpenAI's production.

      RFP 20, 21, 22: For RFP 20, Our reference to "if any" is merely a reference to the fact that such documents may not exist. This is a standard inclusion in many responses to RFPs, and is not substantively different from The Times's limitation in its responses to only produce documents that are in its possession, custody or control located after a reasonable search. The Times's suggestion that documents were improperly destroyed is not well taken. OpenAI has agreed to search for and produce documents within its possession, custody, and control, if any, sufficient to show the total number of records and tokens in OpenAI's training datasets that can be located pursuant to a reasonable and diligent search. If The Times has concerns following OpenAI's production after a reasonable and diligent search, OpenAI will make itself available to discuss those concerns. As to the balance of The Times's request, OpenAI has agreed to make its text training data available for inspection, from which it can identify the information it seeks.

      For RFP 21, OpenAI has trained the relevant models using (a) information that is publicly available on the Internet, (b) agreements related to access to text training data, and (c) other information generated by OpenAI for the purpose of AI training. The agreements are sufficient to identify all sources of information that fall into the second category. OpenAI is also making its training data available for inspection. We do not believe there is any information that The Times needs in addition to this information.

      Finally, for RFPs 21 and 22, OpenAI has responded to requests by making the training data available for inspection because it is in those training datasets where the information sought by the request can be found. You have not explained why documents in addition to the training data or documents OpenAI agreed to produce are insufficient. Given you have not yet engaged in any inspection, any such claim would be premature and unfounded.

      RFP 23: OpenAI did not ignore any portion of the request, and specifically objected to the production of "documents concerning 'all steps . . . contemplated to train, instruct, program, or refine' OpenAI's products" as "not relevant to any claim or defense in this litigation." If The Times disagrees, please be prepared to discuss on any meet and confer why such documents are relevant, and why The Times needs documents concerning "all steps taken or contemplated . . . ." Moreover, please come prepared to clarify the terms "steps" and "contemplated" as these undefined terms make the request overly broad, unduly burdensome and disproportionate to the needs of the case (e.g., "contemplated" by whom? Anybody at OpenAI?). As to the portion addressing Times content or Journalism content, OpenAI has responded to requests by making the training data available for inspection because it is in those training datasets where the information sought by the request can be found, if it exists.

IIIORRISON FOERSTER

August 20, 2024
Page Six

Additionally, OpenAI has agreed to "collect and produce documents related to 'pre-training' and 'post-training,'" which, as OpenAI has previous clarified, is "intend[ed] to cover the entire training process." See June 5 email. As previously confirmed, we understand this to include the "specific aspects of training listed in the RFP." Once you have reviewed that information, we can discuss whether The Times believes information relating to the categories you mention is missing and the basis for that belief.

RFP 26: OpenAI responded to the request by making the training data available for inspection because it is in the training datasets that any data derived from OpenAI's web crawlers and made available to GPT models will be found. We appreciate you explaining in more specificity the documents The Times is actually looking for with this RFP. OpenAI will agree to supplement its response to include the following page on OpenAI's website containing responsive information: https://platform.openai.com/docs/bots. Additionally, it is not apparent to us how "OpenAI's policies and practices for third-party robots.txt directives and/or metatags that block the crawling and scraping of copyrighted content and whether Defendants rely on extending robots.txt permissions surrounding crawling and indexing for Search for GAI purposes" fall within the scope of RFP No. 26 as written. If The Times confirms that those are the documents it wants, we can bring that back to our client to formulate an appropriate response.

RFP 27: As an initial matter, this RFP is facially overbroad. For example, The Times states that this information is "clearly relevant to how ***Defendants*** obtained content to power ***Defendants'*** Generative AI products and Services." (emphasis added). However, facially this RFP calls for information relating to a plethora of products that are not Defendants', including "custom GPTs included on Defendants' platforms, and related applications . . . ." Regardless, OpenAI agreed to search for and produce responsive documents. We are confused as to what The Times believes is deficient with the response. As discussed above, "sufficient to show" limitations are not improper. That said, OpenAI is willing to discuss The Times's objections to better understand what it is seeking that it has not already agreed to produce.

Additionally, we are not clear on what The Times's objection is to OpenAI's limitation to robot.txt. OpenAI's understanding is that all relevant "crawler[s]" are related to robots.txt. *See* https://platform.openai.com/docs/bots.

RFP 28: OpenAI objected to The Times's request for "documents concerning all measures taken to conceal and obscure Defendants' use of Times Content[.]" The Times's use of this language appears to be an attempt to make inferences regarding documents produced that are unwarranted. OpenAI will not adopt this language. Rather than refusing to provide any response as a result of that objection, OpenAI provided a response to the unobjectionable Parts (i), (ii), (iii) and (iv) of the RFP. OpenAI's agreement to produce responsive, non-privileged documents "sufficient to show filtering of Times Content" is directly responsive to those part of the request. By "filtering," OpenAI refers to processes by which it prevents certain output from being provided to a user (e.g., "preventing [ChatGPT]

||||ORRISON ⸫OERSTER

August 20, 2024
Page Seven

from reproducing Training Datasets containing Times Content."). "Filtering" is not a measure taken to "conceal or obscure" Defendants' use of Times Content; OpenAI continues to object to the RFP to the extent it characterizes any alleged actions taken by OpenAI as "measures taken to conceal or obscure Defendants' use of Times Content." OpenAI does not understand why The Times believes this request is considered a "core issue." Please come to the meet and confer so that we can better understand The Times's position and what documents The Times is seeking, and, if necessary, reconsider OpenAI's position.

RFP 31: By "relevant technical documentation," OpenAI simply means the technical documentation for the OpenAI models and products at issue in this litigation. As you may know, Annex IV of the European Union's Artificial Intelligence Act will not apply until 2026. Whether OpenAI's documentation "complies" with a law whose provisions do not yet apply is therefore irrelevant. Regardless, based on OpenAI's investigation to date, its relevant technical documentation can be found at https://openai.com/news/research/. To the extent not already produced, OpenAI will produce the documentation located at that URL.

RFP 32: As discussed above, OpenAI's "sufficient to show" limitation is not improper, particularly given the generality of the documents sought by the RFP. Moreover, the parties have already resolved prior disputes over what "training process" includes. [Levy 6/5 email ("When OpenAI says that it is willing to collect and produce documents related to 'pre-training' and 'post-training,' we are intending to cover the entire training process.")]. OpenAI understands each of the identified categories of information in RFP No. 32 to fall within the training process. Here is an example of a document produced related to "the training process for the GPT models used for ChatGPT": OPCO_NYT_0193164. If The Times believes its production is deficient, please come prepared to the meet and confer to identify those deficiencies.

RFP 36: RFP No. 36 specifically requests "differences in the Training Datasets and/or the composition of the Training Datasets used for each model." OpenAI's response is directly responsive to the request. As to the other portions of the request, OpenAI requests additional clarification. Please come to the meet and confer prepared to explain what kinds of documents are sought by the RFP.

RFP 43: By "OpenAI work," OpenAI refers to any "surveys, studies, analyses, or statistics on patterns and practices" conducted, generated, or commissioned by OpenAI relating to ChatGPT and, per the parties' recent agreement, the ChatGPT API, as requested in RFP No. 43. To the extent such works were prepared by an entity other than OpenAI but is within OpenAI's possession, custody or control, they would be included.

RFP 44: To the extent this information exists and concerns OpenAI's actions, it would be encompassed by OpenAI's response.

RFP 49: First, as explained above, the sufficient to show limitation is not inherently improper. OpenAI will, however, agree to remove its sufficient-to-show limitation as to this

IIIORRISON FOERSTER

August 20, 2024
Page Eight

response and produce documents that show OpenAI's knowledge (or lack thereof), prior to the inception of this lawsuit, of whether NYT content was included in a dataset for any model used for ChatGPT.

Second, OpenAI does not understand why information about copyrighted content other than The Times's content would be relevant to this litigation. (Relatedly, OpenAI is unclear how it could even identify all "Times Content" as defined by the RFPs ("all New York Times interest-specific publications, sub-brands, and products") unless The Times is able to provide OpenAI with a list of all such publications, sub-brands and products). Please come prepared to the meet and confer prepared to explain its relevance. Moreover, this request would appear to require OpenAI to make legal determinations as to the status of particular reference materials to determine whether or not it would be responsive. This sort of request is improper and unduly burdensome.

Lastly, OpenAI limited its response to the period before the lawsuit was filed because the post-lawsuit documents are very likely to all be privileged. We note that the parties agreed that certain categories of documents such as those at issue here "need not be included on a privilege log." (Dkt. 135-1 at 7.) Reviewing these documents which are likely to be privileged, for the potential that some document is not privileged, does not represent a proportionate discovery request.

RFP 50: OpenAI has agreed to produce the policies it maintains that govern the use of its products and provide information about the information that it is able to monitor and collect from the users of its products. OpenAI's understanding is that no documents exist that show how these rules were or were not applied with respect to The Times prior to the inception of the lawsuit. To the extent The Times requests responsive documents after the inception of the lawsuit, those documents are privileged and fall within the categories of documents that the parties agreed "need not be included on a privilege log." (Dkt. 135-1 at 7.)

RFP 60: The Times seeks through its questions answers that are more appropriately addressed via an interrogatory. We also do not understand your reference to the fact that this case "focuses heavily on the 'outputs of Defendants' GenAI models." OpenAI's position is that only the models used for ChatGPT are relevant to this litigation. As discussed above, if the Court rules in The Times's favor with respect to what models are at issue, OpenAI will of course evaluate how and whether the order applies to these RFPs.

Additionally, OpenAI does not understand The Times's assertion that OpenAI's response to this RFP is inconsistent with OpenAI's agreement to "produce RAG-related documents." Please come to the meet and confer prepared to explain The Times's position.

RFP 62: OpenAI has produced over a dozen agreements to date, including all agreements that were in effect prior to the litigation and therefore relate to text training data used for the models used for ChatGPT at issue in this case. OpenAI is in the process of a

ΙΙΙORRISON ΞOERSTER

August 20, 2024
Page Nine

supplemental collection relating to agreements that were entered into more recently. We cannot give you a timeline as to when that production will be complete. Regardless, OpenAI can confirm that it has been working on securing permission from the relevant third parties to produce the more recent agreements.

Additionally, if the Court rules in Plaintiff's favor with respect to what models are at issue, OpenAI will of course evaluate how and whether the order applies to these RFPs.

RFP 70: The sufficient to show limitation is not improper, as discussed above. OpenAI does not understand how this RFP could be considered a "core issue" to the litigation such as that in *StubHub*. If The Times disagrees, please come to the meet and confer prepared to explain that to us.

OpenAI directs The Times to OPCO_NYT_0192848. If The Times desires additional financial information, please come to the meet and confer prepared to explain why the financial information provided is inadequate.

RFP 71: OpenAI is not excluding profit from its response and directs The Times to the profit and loss statement OpenAI has produced.

Respectfully,

/s/ Allyson R. Bennett
Allyson R. Bennett
Morrison & Foerster LLP